Event-Driven Finance – Course Package

FRE GY9743

http://www.experimental-finance.com

Mike Lipkin mdl9606@nyu.edu
Sacha Stanton as15139@nyu.edu

Table of Contents

Introduction

Please refer to the Laboratory Package for details on the course and access to the data systems.

PROBLEM SET 1 – An introduction to SQL

Explore the tables and return some data for security, security_price, option_price. Get familiar with the columns and looking at the data. You can use SELECT TOP 100 * FROM <table>. (This will return only the first 100 rows. **Do not forget this down the road since you will otherwise be returning all data in the database.**)

1) How many unique tickers are listed in IVY?
2) Pick a day: how many unique tickers trade?

   NOTE: Textbooks tend to give exact answers. In the world of real securities, ambiguities are commonplace, and data analysis is always subtly marred by noise, duplication and fuzzy answers.

3) Select all options on a particular day for an optionable stock (i.e. for which options are listed) beginning with the first letter of your last name.
4) Count all stocks trading at the beginning of each odd year for all years in the database.

   NOTE: Take a look at issue types for stocks – what do they mean? How do they affect the results? This will be important for your projects.

5) What is the (sematic) difference between the two queries below? Explain clearly with no more than 3-4 lines.

   (HINT: Don't pontificate. Run it!)

Query1:

```
SELECT    s.*, sp.*
FROM      security s
LEFT JOIN security_price sp
ON        s.securityID=sp.securityID
AND       sp.date='2020-01-13'
WHERE     sp.closePrice is null
```

Query2:

```
SELECT    ticker
FROM      security s
LEFT JOIN security_price sp
ON        s.securityID=sp.securityID
WHERE     sp.closePrice is null AND
          sp.date='2020-01-13'
```

6) Strike values are stored as integers in IVY. They correspond to a multiple of the real strike of the option, i.e. IVY-strike=1000*real-strike. Create a query that returns the real dollar value of an option strike. Warning: don't make rounding errors!

7) a) plot the (adjusted) price of Tesla over the year 2020. Adjusted means that you need to account for splits;

   b) plot the histogram of returns;

   c) find those days for which the stock return exceeds (in absolute terms) 2.5%.

8) Select the minimum and maximum prices by month, for a stock of your choosing.

9) For each day in a month of your choosing, find the at-the-money (ATM) strike for a particular stock. How many values does it take?

   NOTE: are you confused by this question? What does at-the-money mean? Give it your best guess – we will be discussing at least three meanings for the term, whose relevance depends on the financial context. Even though at-the-money is a fuzzy concept, it is one that is relevant all the time and whose subtleties are important to understand.

10) Consider the company PG:

   a) Plot the implied volatility surface on 03/04/2020. Avoid bad values, i.e. IV<0.

   b) For the period Feb 2020 to Feb 2021, find the IV of the ATM (at the money) front month (i.e. with closest expiration) option and plot it against time. Again, avoid bad values.

11) Count the number of option prices in a one-month period, grouped by week.

12) Find all non-optionable stocks. Reminder: not every security in IVY is a stock.

13) How many tickers have been associated with more than one company?

14) How many companies have had multiple tickers?

PROBLEM SET 2 – Introduction to SQL in Finance

NOTE: Problem 6 is an important and lengthy problem that may take as long as 1-5 combined, so give yourself enough time.

1) Option prices in IVY are presented as *longs*. Write a SELECT statement that returns prices rounded and truncated to the nearest penny.

2) Create a SELECT statement producing the midpoint best bid and offer price (**MBBO**) for every option. This is the mean value of the highest bid and lowest offer from among all exchanges for any particular option. In IVY, the prices are already the national best bid best offer of all exchanges, but in actuality different exchanges have or can have different BB or BOs.

   NOTE: MBBO is an important concept and frequently used value. Ensure you understand what it means and how it is calculated.

3) Create a SELECT statement that rounds the MBBO to the nearest penny.

4) At closing, and even during the course of the day, far out-of-the-money options will have stray bids and offers which make less meaningful the MBBO. The following might be a typical closing set of quotes:

XYZ close 40.21		
Strike	**Calls**	**Puts**
20		0.00 0.10
22.5		0.00 0.05
25		0.05 0.05
…		
50	.10 .15	
55	.05 .25	
60	.00 .05	

   Here just the deep out-of-the-money puts and calls are shown. As you can see, the mbbo for the puts is 0.05, <u>0.025</u>, 0.05 and for the calls is .125, <u>.15</u>, .025. This is obviously not monotonic.

   Note: Price flips only occur for same expiration, same date, and same putCall value.

   Take the stock KO for the six month period 3/1/20 to 9/1/20 and find all the closing price flips of this kind in the mbbo, where the offer is less than $0.10.
   You can use the functions dbo.formatStrike() and dbo.mbbo(bid, ask).

5) Bad data appears from time to time in all fields. For TSLA, from 2019-2020, find the percentage of bad implied volatilities in the option_price_view table.

   NOTE: What do you mean by bad data? What are your criteria?

6) **Alert: Hard Problem!**

   IVY provides a table of yield curves for any given date. Ideally this would fit with put-call parity for any option series. In fact, it does not. What are four possible reasons why put-call parity may fail (in the context of American options this means that implied volatilities are *apparently* different for the puts and calls on any given line)?

   NOTE**:** These are real financial events or reasons, not mathematical peculiarities.

   Take two stocks, one paying no dividend, one paying a dividend. Assuming equal implied volatilities for the MBBO's of the 50-delta puts and calls, construct two approximate yield curves for both options on the dates: 2/13/20, 4/16/20 and 7/10/20. Compare your interest rates to the yield curves in IVY.

   For this problem you can use VBA, Matlab or Python and must use an American option pricer (one for Excel can be downloaded from the class website).

   NOTE**:** terms such as "50-delta" or "at-the-money" are heuristic, intended to identify the strike closest to the stock price.

   WARNING: there is no exact "50-delta" option in practice.

PROBLEM SET 3 – Synthetic Structures

Pick one of the following three stocks: CSCO, TSLA, IBM.

1) For the 2nd series (2nd month) over the one year period of 2019, find all puts with more than $0.50 Premium Over Parity (POP). Of these, rank the volatility of each strike versus its moneyness (defined here as ratio of strike price to stock price).

   For normally skewed options the volatility should be monotonically declining with moneyness. Are there any notable exceptions?

   NOTE: Theorists but not practitioners refer to POP as the "extrinsic value."

2) Consider the day-to-day volatility of the 2nd month ATM option. Find the three biggest overnight absolute changes in 2020. List the dates and attempt to determine the cause online.

   NOTE: make sure you define your interpretation of ATM.

3) Sometimes you may want to follow an option that does not actually trade. Synthetic options are used as proxies, and derived by interpolating data from nearby options that actually trade.

   For each day in a week of your choosing, construct "synthetic ATM 45-day" put and call options with the following properties:

   a) the strike is constructed to be exactly the closing price
   b) the expiration date is 45 days ahead

   Find the implied volatilities by inverting Black-Scholes for the interpolated prices. Repeat this for two additional series with strikes located +/- 10% of the synthetic ATM.

   NOTE: to validate your synthetics, ensure that implied volatilities of the puts and calls are approximately equal.

4) Go to an online charting tool (e.g. Google finance) and look at the (current) one-year plot of KO vs. PEP.

   Suppose this motivates you to pairs trade. (One would still be curious whether the best strategy is to trade stock price or volatility.)

   Create a similar plot for the synthetic 45-day ATM option volatilities for both stocks for a three-year period starting 2018.

   At first glance, does the ratio seem to mean revert?

PROBLEM SET 4 – Pinning

NOTE: This problem set will **take time.** Start early, and coordinate with your fellow students on when you will use the server.

Expirations are listed in IVY as the day after the option last trades. E.g. if an option expires at 4pm on Friday, it will be listed in IVY as Saturday. Bear this in mind when constructing your queries.

1) **Warning: Long Problem!**
   Consider INTC, KO, and then all optionable stocks in IVY (not indices, see the index flag on the security table). For the period of 1/1/1996 to 6/1/2021 for INTC and KO, reproduce the PPN bar graph for T-5 to T+5 trading days about expirations using a cut-off of $0.15 to define pinning. Do the same for all optionable stocks for the period 1/1/2020 to 1/1/2021. How reasonable are the statistics for these abbreviated sets (especially single stocks)?

   NOTE: This problem involves significant data crunching so may take a while to run – bear in mind however that not observing index optimization will cause it to never finish. Try a reduced range at first.

   NOTE: You are only interested in trading days - don't get caught by weekends/holidays.

2) Avellaneda-Lipkin predicts monotonicity of pinning probability in the parameter N/σ/(avg. daily stock volume). For all the stocks in the S&P 500 from 01/01/2019 to 01/01/2021:

   Recreate the PPN graphs, binning the ATM strikes into quartiles by open interest/(avg. daily stock volume). You will want to take:

   - Open Interest on expiration day
   - Avg. daily stock volume calculated over 5 trading days before expiration

   For the entire market over the same time period, recreate the PPN graphs, binning the ATM strikes into quartiles by open interest/(avg. ATM implied vol.)/(avg. daily stock volume). You will want to take:

   - Avg. ATM implied volatility calculated over 5 trading days before expiration

   [WARNING: Again, this problem might take a while to compute. It would be wise to retain the data for use in problem 3.]

3) Avellaneda-Lipkin also predicts that the probability of stock pinning should fall away with (logarithmic) distance from the strike. For the S&P 500 data set already collected in problem 2, find a) the logarithmic distance in price to the nearest strike one week prior to expiration for every expiration, b) whether the stock pinned.

Aggregate all the data and plot the probability of pinning vs. logarithmic distance.

4) ETFs and Indices. Previous classes showed that indices DID NOT PIN. Now there are many ETFs that may be traded in lieu of the individual names. Pick 3 unrelated ETFs and see how often they pin in a three-year period of your choosing.

PROBLEM SET 5 – Dynamics

1) Pick three optionable stocks.

   a) Using the Internet, make a table of announced earnings dates for the two-year period 6/1/2019-6/1/2021.

   b) For each stock, identify the option series that will be: (A) the front month at earnings, (B) the next available series, (C) the first January leap (this will be no sooner than the fifth available option month).

      NOTE: For example, if earnings are Feb. 6, the series in question are Feb, Mar, and then the subsequent Jan. If earnings are Feb. 25 the series in question are Mar, Apr, and then the subsequent Jan.

2) We wish to follow the implied volatilities of the ATM straddles of various series beginning approximately three weeks before earnings and proceeding to one week after earnings.

   NOTE: for this problem, we do not use the volatility surface tables in IVY to track the 50-delta vols. Why, you ask? The volatility surface table uses time averaging and therefore minimizes drops across earnings events.

   a)
      - Choose one of the stocks from 1a.
      - For each earnings event, create three synthetic ATM options, one for each of the three series identified in 1b
      - plot their implied vols for the 4 week period

      The result will be 3 series x 4 earnings = 12 curves.

      NOTE: The front month series may truncate with expiry but the other two will continue past expiration.

   b) Do volatilities drop discontinuously across the earnings dates?

   c) Try to fit the volatility profiles running up to earnings with parabolas. Can you see any regularity? Would exponentials do better?

   d) What difficulty arises with following individual strikes that we avoid by following synthetic strikes?

3) Recall in the first lecture we discussed the possible impact of a large trade. Having a minute database allows us to examine the consequences of such a trade on a finer scale.

   Pick a stock from the LiveVol database and find the largest options trade to occur at least one week removed from an Earnings date and at least two times larger than any trade for the subsequent 3 days. We wish to examine minute by minute over the 3-day period.

   a) Generate an implied volatility surface for your choice just prior to the large trade. If space is an issue concentrate on a surface centered about the strike and series of the large trade

   b) Follow the vol surface minute by minute after the large trade

   c) Can you characterize a relaxation time scale for the vol surface

   d) Propose a possible trading scheme for a high frequency trader to monetize a future disturbance of this kind

PROBLEM SET 6 – Hard-To-Borrows

1) Take the hard-to-borrow stock GME over the period 2020 through 2021. Look at a synthetic 2.5-month series (one always 75 days ahead), both puts and calls.

   Ignoring early exercise (is this a good assumption?), calculate the *negative* interest rates that produce put-call parity between MBBO's and plot these.

   NOTE: Remember your use of the pricer in Problem Set 2. Vary the interest rate until the implied vols for the puts and calls become equal.

2) We can redo (1) more accurately. Lipkin-Avellaneda predicts an effective dividend rate (term structure) restores put-call parity in HTBs. For the 2-year period above, calculate the entire term-structure of dividends (4 months + 2 leaps) on the first trading day of each month. You will need the American options pricer to account for early-exercise premia and the zero curve table to get the ordinary interest rates.

   How do the 3-month effective dividend rates compare with the negative interest rates found in 1)?

   NOTE: You are plotting dividend rate vs. time to expiration.

3) Redo problems 1 and 2 with the stock SAVA for the same period.

PROBLEM SET 7 – Take-Overs

1) Find two optionable stocks which were acquired in the period 1/1/2012 to 1/1/2021, one for cash and one for stock. Make note of the initial announcement dates.

   Caution: an announcement date does not guarantee an acquisition eventually occurs.

2) Examine the option series for both stocks from 3 months before to one month after the *announcement* of the deal. We want to concentrate on the “50-delta” (ATM) and “40-delta” (just out-of-the-money) options in the front months and the “30-delta” (more out-of-the-money) options in the back months.

   a) Do the option prices (implied volatilities) conform to the “caricature” discussed in class?

   b) Were these deals *leaked/anticipated*? If so, how much lag time did the “good guessers” have?

3) For the stock CEPH, look at the period of 01/2011 to 05/2011.

   a) What significant event seems to happen? When?

   b) The term structure of volatility refers to the variation of implied volatility with respect to maturity date. Plot the “30/40/50-delta” implied volatilities with respect to expiration from 01/2011 through 05/2011. What, if anything, do you see?